



Open + Reproducible Research Webinar for USGS

Practical steps for increasing the
openness and reproducibility of
research

April Clyburne-Sherin

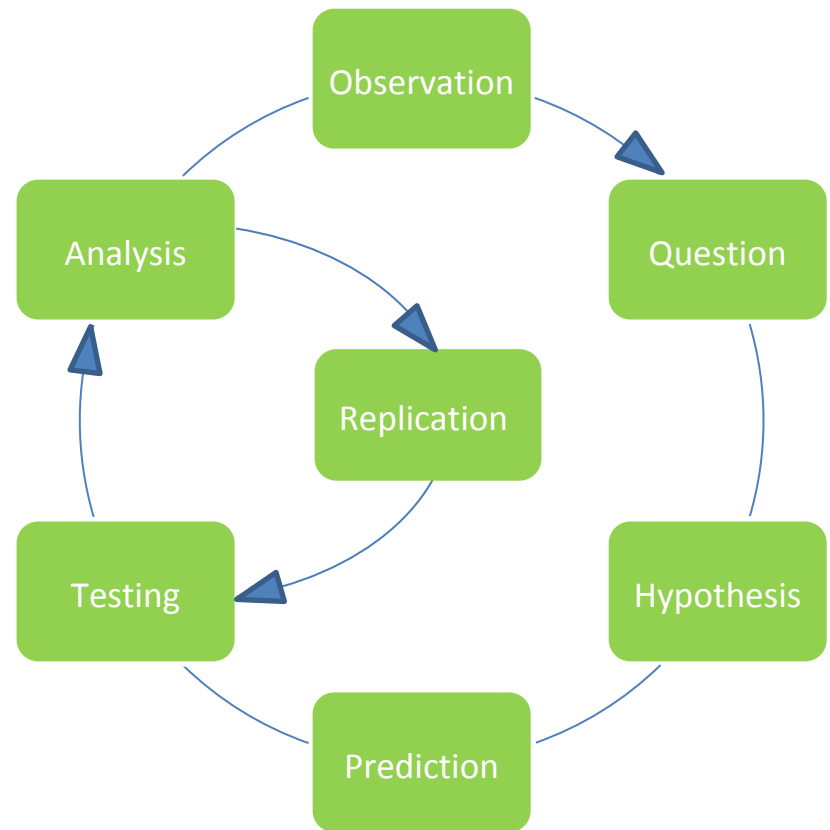
Objectives

- Understanding reproducible research
- Setting up a reproducible project
- Preregistering
- Keeping track of things
- Containing bias
- Sharing

What is the problem?

Scientific method

- Best way to learn how the world works
- Replication of findings is highest standard of evaluating evidence
- Replication of methods allows reuse and extension of new knowledge



What is the problem?

CORRESPONDENCE

[LINK TO ORIGINAL ARTICLE](#)

Open access, freely available online

Believe it or not: how much can we rely on published data on potential drug targets?

Florian Prinz, Thomas Schlange and Khusru Asadullah

A recent report by Arrowsmith noted that the success rates for new development projects in Phase II trials have fallen from 28% to 18% in recent years, with insufficient efficacy being the most frequent reason for failure (Phase II failures: 2008–2010. *Nature Rev. Drug Discov.* 10, 328–329 (2011))¹. This indicates the limitations of the predictivity of disease models and also that the validity of the targets being investigated is frequently questionable, which is a crucial issue to address if success rates in clinical trials are to be improved.

to 'feasible/marketable', and the financial cost of pursuing a full-blown drug discovery development programme for a particular target could ultimately be hundreds of millions Euros. Even in the earlier stages, investment in activities such as high-throughput screening programmes are substantial, and thus the validity of published data on potential targets is crucial for companies when deciding to start novel projects.

To mitigate some of the risks of such investments ultimately being wasted, most phar-

Power failure: why small sample size undermines the reliability of neuroscience

Katherine S. Button^{1,2}, John P. A. Ioannidis³, Claire Mokrysz¹, Brian A. Nosek⁴, Jonathan Flint⁵, Emma S. J. Robinson⁶ and Marcus R. Munafò¹

Abstract | A study with low statistical power has a reduced chance of detecting a true effect, but it is less well appreciated that low power also reduces the likelihood that a statistically significant result reflects a true effect. Here, we show that the average statistical power of studies in the neurosciences is very low. The consequences of this include overestimates of effect size and low reproducibility of results. There are also ethical dimensions to this problem, as unreliable research is inefficient and wasteful. Improving reproducibility in neuroscience is a key priority and requires attention to well-established but often ignored methodological principles.

Essay

Why Most Published Research Findings Are False

John P.A. Ioannidis

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance.

factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings

Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenient, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a p -value less than 0.05. Research is not most appropriately represented and summarized by p -values, but, unfortunately, there is a widespread notion that medical research articles

It can be proven that most claimed research findings are false.

should be interpreted based only on p -values. Research findings are defined here as any relationship reaching formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. "Negative" research is also very useful. "Negative" is actually a misnomer, and the misinterpretation is widespread. However, here we will target relationships that investigators claim exist, rather than null findings.

As has been shown previously, the probability that a research finding is indeed true depends on the prior probability of it being true (before doing the study), the statistical power of the study, and the level of statistical significance [10,11]. Consider a 2×2 table in which research findings are compared against the gold standard of true relationships in a scientific field. In a research field both true and false hypotheses can be made about the presence of relationships. Let R be the ratio of the number of "true relationships" to "no relationships" among those tested in the field. R

is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is $R/(R+1)$. The probability of a study finding a true relationship reflects the power $1 - \beta$ (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate, α . Assuming that c relationships are being probed in the field, the expected values of the 2×2 table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, the post-study probability that it is true is the positive predictive value, PPV. The PPV is also the complementary probability of what Wacholder et al. have called the false positive report probability [10]. According to the 2×2 table, one gets $PPV = (1 - \beta)R / (R - \beta R + \alpha)$. A research finding is thus

Citation: Ioannidis JPA (2005) Why most published research findings are false. *PLoS Med* 2(8):e124.

Copyright: © 2005 John P.A. Ioannidis. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abbreviation: PPV, positive predictive value

John P.A. Ioannidis is in the Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, and Institute for Clinical Research and Health Policy Studies, Department of Medicine, Tufts New England Medical Center, Tufts University School of Medicine, Boston, Massachusetts, United States of America. E-mail: ioannidis@ccuoi.gr

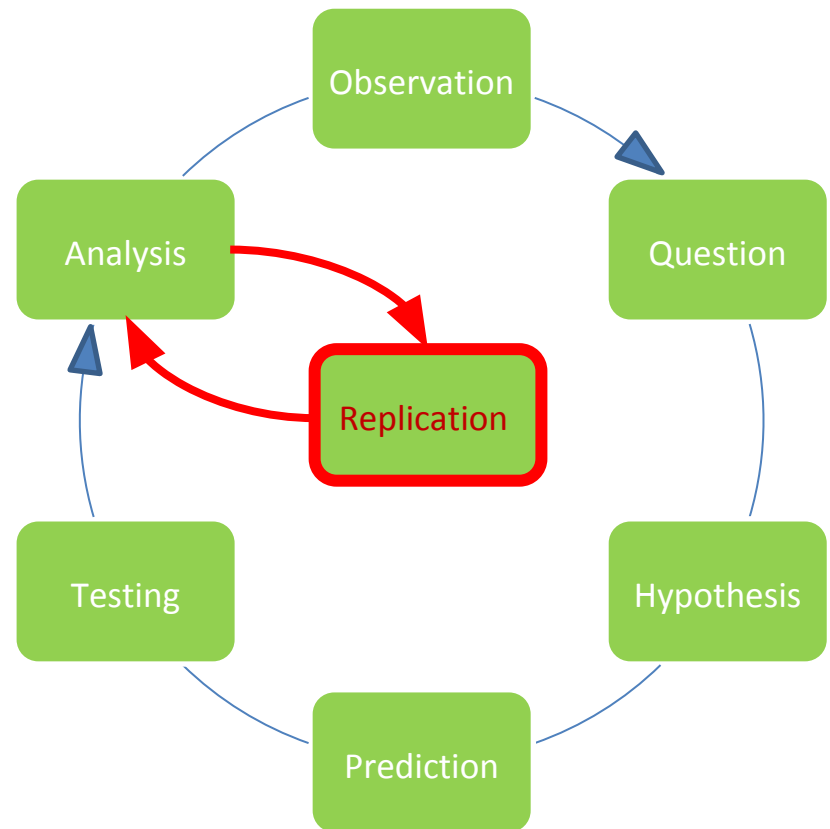
Competing Interests: The author has declared that no competing interests exist.

DOI: 10.1371/journal.pmed.0020124

What is reproducibility?

- **Computational reproducibility**

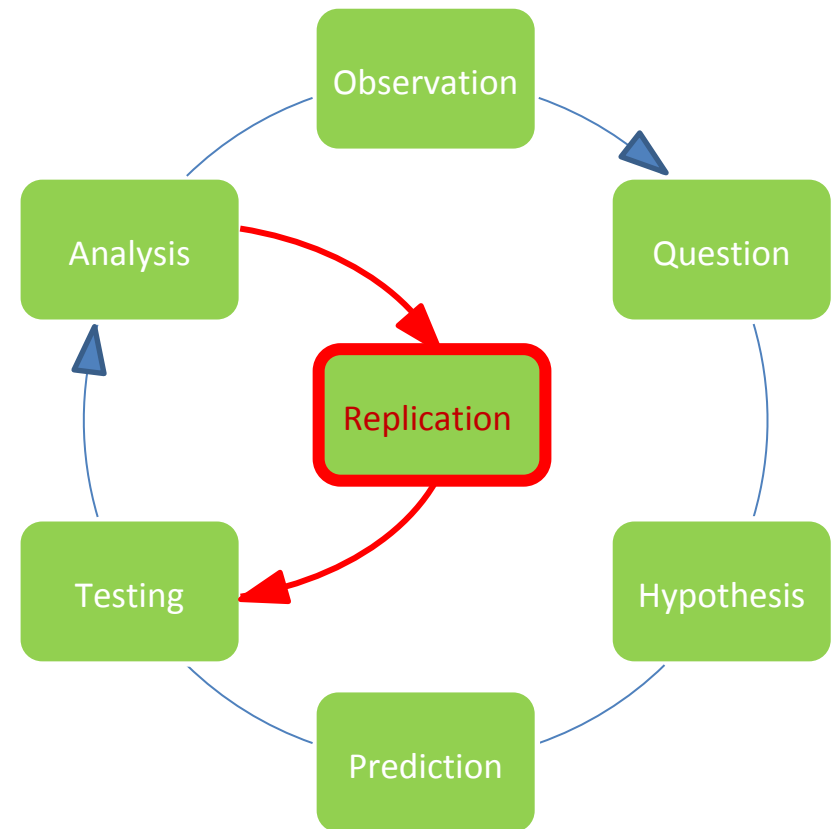
Scientific method



What is reproducibility?

- Computational reproducibility
- **Empirical reproducibility**

Scientific method

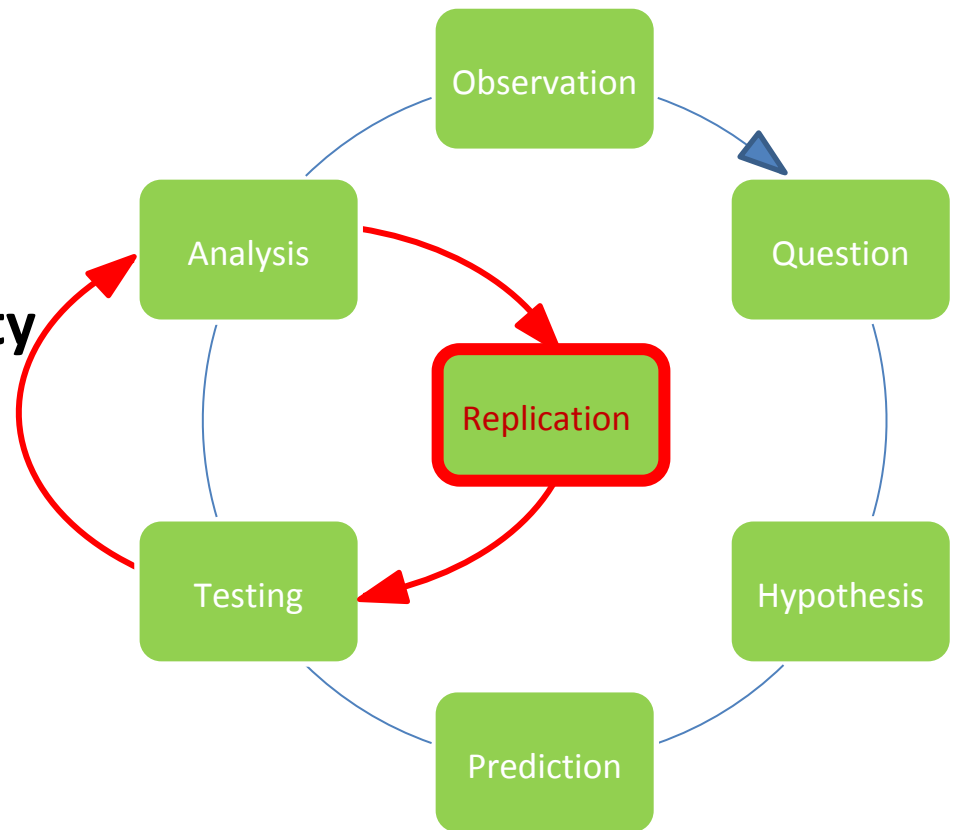


What is reproducibility?

- Computational reproducibility
- Empirical reproducibility
- **Conceptual reproducibility**

New data

Scientific method



What are the barriers?



What are the barriers?

- Statistical
 - Low power
 - Researcher degrees of freedom
- Transparency
 - Poor documentation
 - Poor reporting
 - Lack of sharing

Why practice reproducibility?

The idealist

- **Shoulders of giants!**
- Validates scientific knowledge
- Allows others to build on your findings
- Improved transparency
- Increased transfer of knowledge
- Increased utility of your data + methods

The pragmatist

- Increased efficiency
- Reduces false leads based on irreproducible findings
- Data sharing citation advantage (Piwowar 2013)
- “It takes some effort to organize your research to be reproducible... the **principal beneficiary is generally the author herself.**” – Schwab & Claerbout

How can you make your research reproducible?

1. Plan for reproducibility before you start

- Create a study plan
- Set-up a reproducible project
- Preregistration
- Informative file naming

2. Keep track of things

- Version control
- Documentation

3. Contain bias

- Reporting

4. Archive + share your materials

1. Plan for reproducibility before you start

Create a study plan

- **Create a study plan** before you gather your data
- Create a **data management plan**
- Begins documentation early
- Shows evolution of study

How?

- Research questions
- Study design
 - Type of design
 - Sampling
 - Power and sample size
 - Randomization?
- Variables measured
 - Meaningful effect size
- Variables constructed
 - Data processing
- Data management plan
- Analysis plan
- Sharing plan

1. Plan for reproducibility before you start


Set-up a reproducible project





How?

- **Set-up a centralized location for project management** <https://osf.io/>
- Organization is especially important for collaboration
- Easily find the most recent file version
- Eases transition between lab members
- Allows for backup and version control





1. Create an OSF project

 Open Science Framework

My DashboardExplore ▾Help ▾ Courtney Soderberg

Projects

[Add Collection](#) [Add Existing Project](#) [Search](#) 

Name ^ ▾	Contributors	Modified
 Dashboard	Soderberg	3 days ago, by Soderberg
● All my registrations		
● All my projects		

Quick Tasks

Watchlist

Go to my project

You do not have any projects yet. Click below to create one!

Create a project

Title

Description

Template (Optional)
Start typing to search. Selecting project as template will duplicate its structure in the new project without importing the content of that project.

Create

Register a project

Upload file(s)

1. Wiki, file tree, components, citation, GUID

Demo Project

FilesWikiAnalyticsRegistrationsForksContributorsSettings

Demo Project

PrivateMake Public+00

Contributors: Courtney Soderberg

Date created: 2015-12-10 10:08 AM | Last Updated: 2015-12-10 10:08 AM

Category: Project

Description: A project created to learn how to use the OSF.

License: No license

Wiki

No wiki content

Files

Search

Name ^ v

Project: Demo Project

OSF Storage

Citation

osf.io/pg563

Components

Add ComponentAdd Links

No components have been added to this project.

Tags

add a tag

Recent activity

All times displayed at -0500 UTC offset.

2015-12-10 10:08 AM Courtney Soderberg created Demo Project

1. Giving contributors access

[Demo Project](#) [Files](#) [Wiki](#) [Analytics](#) [Registrations](#) [Forks](#) **Contributors** [Settings](#)

Permissions ?


Administrator
Read + Write
Read

Bibliographic Contributor ?

Bibliographic
Non-Bibliographic

Contributors [+ Add](#)

Drag and drop contributors to change listing order.

Name	Permissions ?	Bibliographic Contributor ?	
 Courtney Soderberg	<div>Administrator</div>	<input checked="" type="checkbox"/>	Remove

View-only Links [+ Add](#)

Create a link to share this project so those who have the link can view—but not edit—the project.

1. Creating a wiki

[Demo Project](#) [Files](#) [Wiki](#) [Analytics](#) [Registrations](#) [Forks](#) [Contributors](#) [Settings](#)

[Home](#)

[+ New](#)

Project Wiki Pages

Home

View

Preview

Research Question


















The purpose of this project is to explore what factors are associated with people's beliefs that they can influence government policy.

Hypothesis

This research is exploratory, so we have no a priori hypotheses.

Edit

Live editing mode

B *I*                 

****Research Question****

The purpose of this project is to explore what factors are associated with people's beliefs that they can influence government policy.

****Hypothesis****

This research is exploratory, so we have no a priori hypotheses.

Revert

Save

1. Adding organizational structure - components

Add component ×

Data

Data

☐ Add contributors from **Demo Project**

Cancel

Add

How can you make your research reproducible?

1. Plan for reproducibility before you start

- Create a study plan
- Set-up a reproducible project
- Preregistration
- Informative file naming

2. Keep track of things

- Version control
- Documentation

3. Contain bias

- Reporting

4. Archive + share your materials

1. Plan for reproducibility before you start

Preregister your study plan

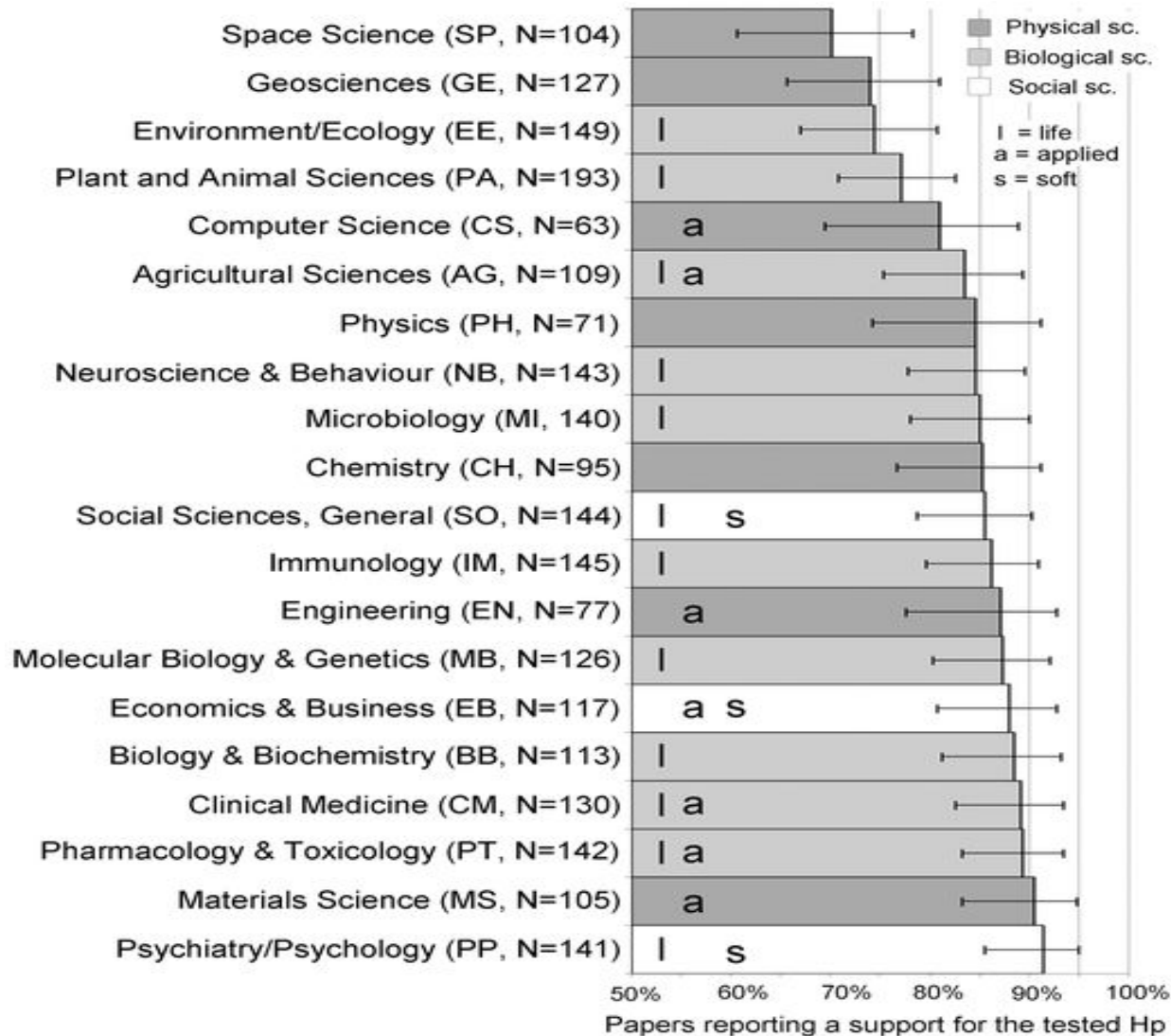
- **Preregister your study plan** before you look at your data
- Distinguishes *a priori* design decisions from *post hoc*
- Counters selective reporting and outcome reporting bias
- Preregistration of all study plans helps counter publication bias

PROSPERO

ClinicalTrials.gov

AEA RCT Registry

Publication bias



1. Plan for reproducibility before you start

Analysis plan

- **Preregister your analysis plan** before you look at your data
- Defines your confirmatory analyses
- Decreases researcher degrees of freedom

How?

- Define data analysis set
- Statistical analyses
 - Primary
 - Secondary
 - Exploratory
- Missing data
- Outliers
- Multiplicity
- Subgroups + covariates

(Adams-Huet and Ahn, 2009)

Researcher degrees of freedom

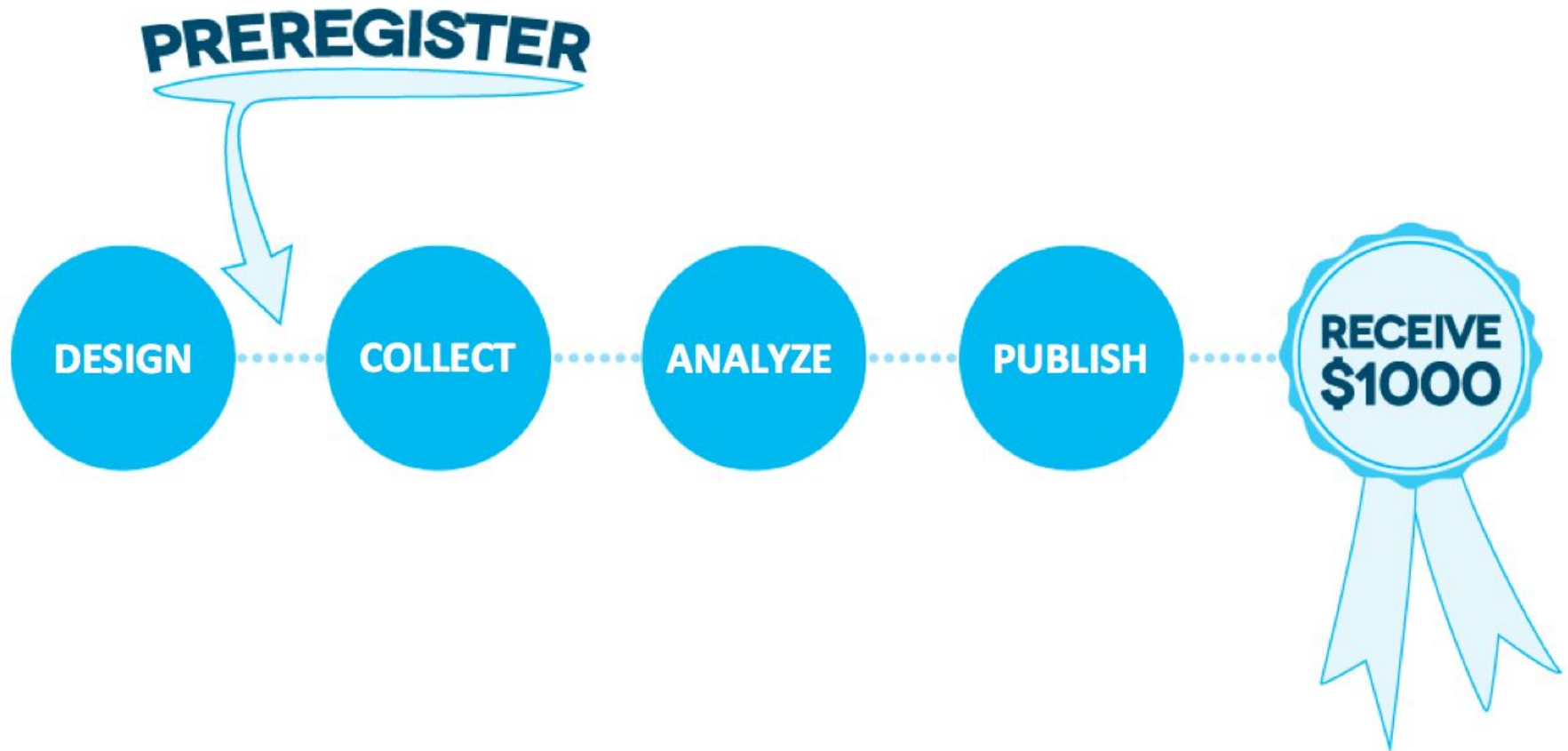
Table 1. Likelihood of Obtaining a False-Positive Result

Researcher degrees of freedom	Significance level		
	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ($r = .50$)	17.8%	9.5%	2.2%
Situation B: addition of 10 more observations per cell	14.5%	7.7%	1.6%
Situation C: controlling for gender or interaction of gender with treatment	21.6%	11.7%	2.7%
Situation D: dropping (or not dropping) one of three conditions	23.2%	12.6%	2.8%
Combine Situations A and B	26.0%	14.4%	3.3%
Combine Situations A, B, and C	50.9%	30.9%	8.4%
Combine Situations A, B, C, and D	81.5%	60.7%	21.5%

1. How to preregister

The screenshot displays the Open Science Framework (OSF) interface. At the top, the header includes the OSF logo, the text 'Open Science Framework', and navigation links: 'My Dashboard', 'Browse', 'Help', a search icon, a user profile for 'April Clyburne-Sherin', and settings and share icons. Below the header, a sub-header shows the project title 'An island at risk: geographic epidemiolo...' and links for 'Files', 'Wiki', 'Analytics', 'Registrations' (which is highlighted in blue), 'Forks', 'Contributors', and 'Settings'. The main content area features a 'Registrations' tab and a message: 'There have been no completed registrations of this project. For a list of the most viewed and most recent public registrations on the Open Science Framework, click [here](#). You can start a new registration by clicking the "New registration" button, and you have the option of saving as a draft registration before submission.' A green button labeled 'New registration' is positioned to the right of the message.

The Preregistration Challenge



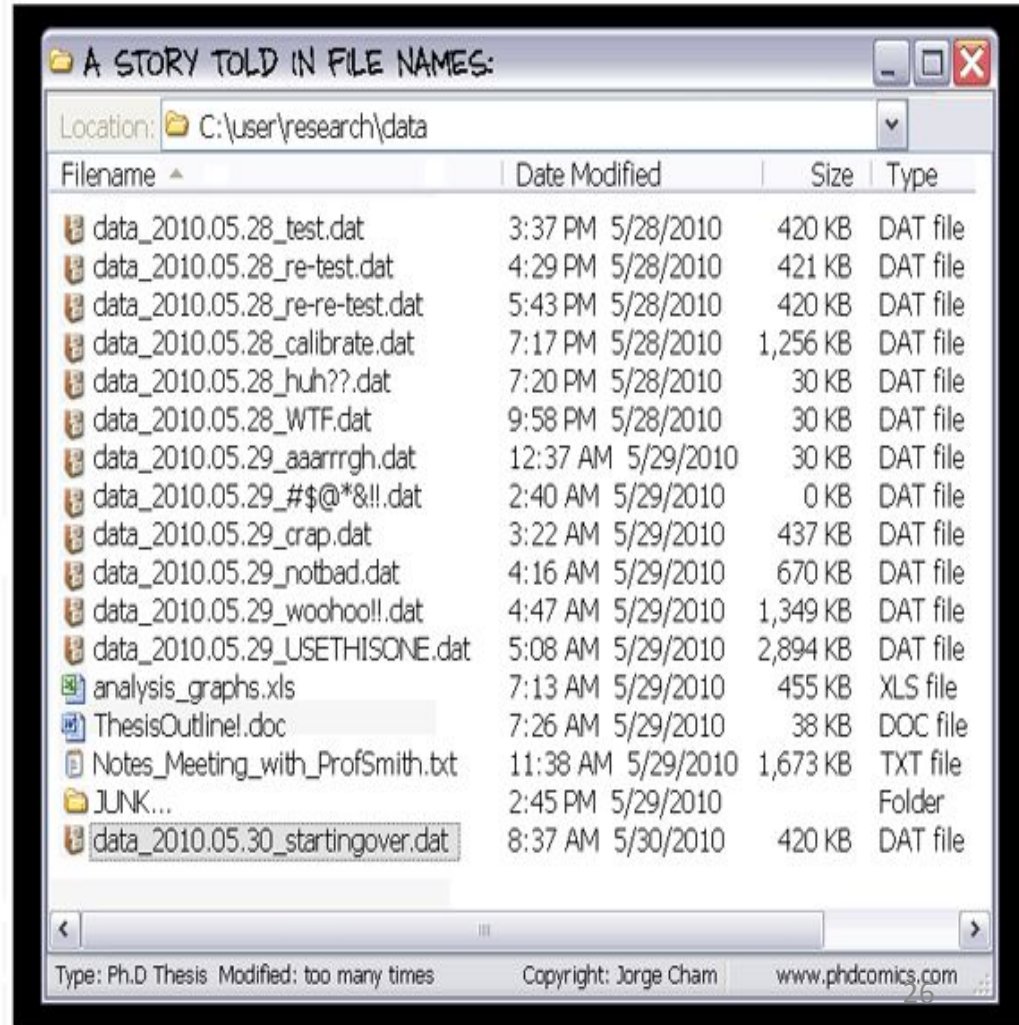
\$1000 incentive to
preregister

<http://cos.io/prereg>

1. Plan for reproducibility before you start

Informative name + location

- Plan your file naming + location system *a priori*
- Names and locations should be distinctive, consistent, and informative:
 - What it is
 - Why it exists
 - How it relates to other files



1. Plan for reproducibility before you start

Informative name + location

- **The rules don't matter. That you have rules matters.**
- Make it machine readable:
 - Default ordering
 - Use of meaningful delimiters and tags
 - Example: use “_” and “-” to store metadata in name (eg, YYYY-MM-DD_assay_sample-set_well)
- Make it human readable:
 - Choose self-explanatory names and locations

2. Keep track of things

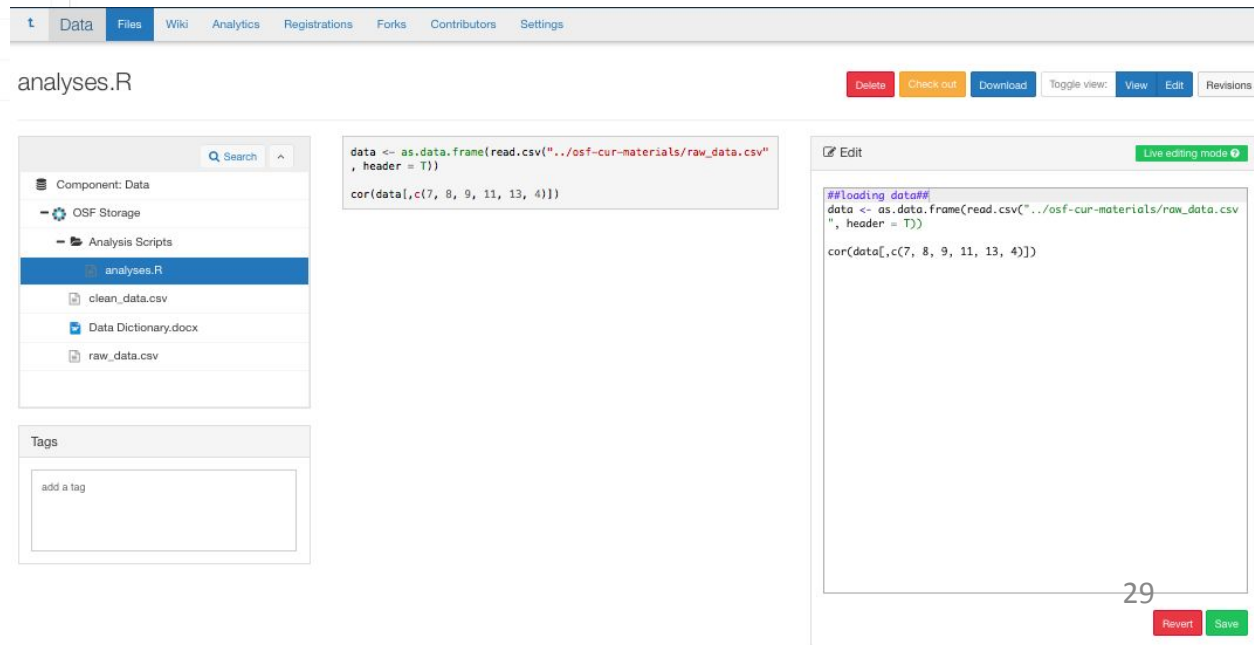
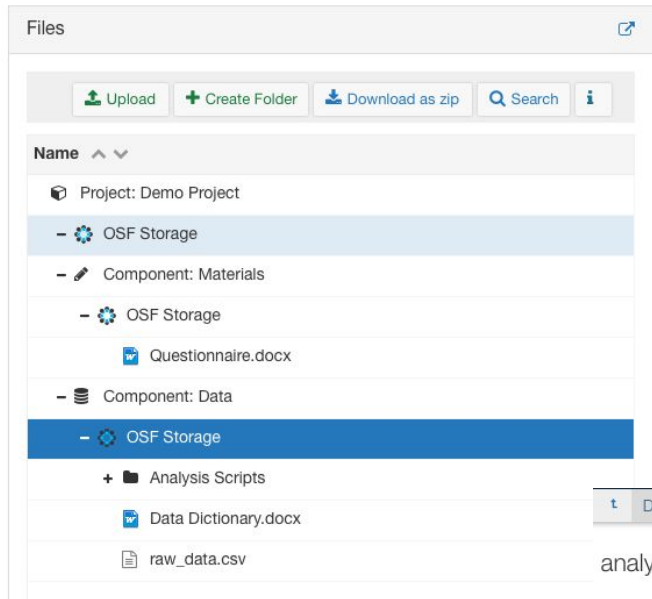
Version control

- **Track your changes**
- Everything created manually should use version control
- Tracks changes to files, code, metadata
- Allows you to revert to old versions
- Make incremental changes: commit early, commit often
- Git / GitHub / BitBucket

Version control for data

- Metadata should be version controlled

2. Version control



2. Version control

[Materials](#) [Files](#) [Wiki](#) [Analytics](#) [Registrations](#) [Forks](#) [Contributors](#) [Settings](#)

Questionnaire.docx [Delete](#) [Check out](#) [Download](#) [View](#) [Revisions](#)

Search

Component: Materials

OSF Storage

Questionnaire.docx

Tags

add a tag

Version ID	Date	User	Download	MD5	SHA2
2	2015-12-11 02:54 PM	Courtney Soderberg	0	6965c573505b9c7503	68bdf73db9cad44fac8
1	2015-12-11 02:50 PM	Courtney Soderberg	1	29cc1757b39a98bda6	978b127c88306cd1ca

2. Keep track of things

Documentation

- **Document everything done by hand**
- Document your software environment (eg, dependencies, libraries, `sessionInfo ()` in R)
- Everything done by hand or not automated from data and code should be precisely documented:
 - README files
- Make raw data read only
 - You won't edit it by accident
 - Forces you to document or code data processing
- Document in code comments

1. Add-ons

Select Add-ons

Documentation

☒ Wiki

Storage

☐ Amazon S3

☐ Box

☐ Dataverse

☐ Dropbox

☐ figshare

☐ GitHub

☐ Google Drive

☒ OSF Storage

Other

☐ External Link

Citations

☐ Mendeley

☐ Zotero

Apply

3. Contain bias



Reporting

- **Report transparently + completely**
- Transparently means:
 - Readers can use the findings
 - Replication is possible
 - Users are not misled
 - Findings can be pooled in meta-analyses
- Completely means:
 - All results are reported, no matter their direction or statistical significance

How?

- Avoid HARKing:
Hypothesizing After the Results are Known
- Report all deviations from your study plan
- Report which decisions were made after looking at the data

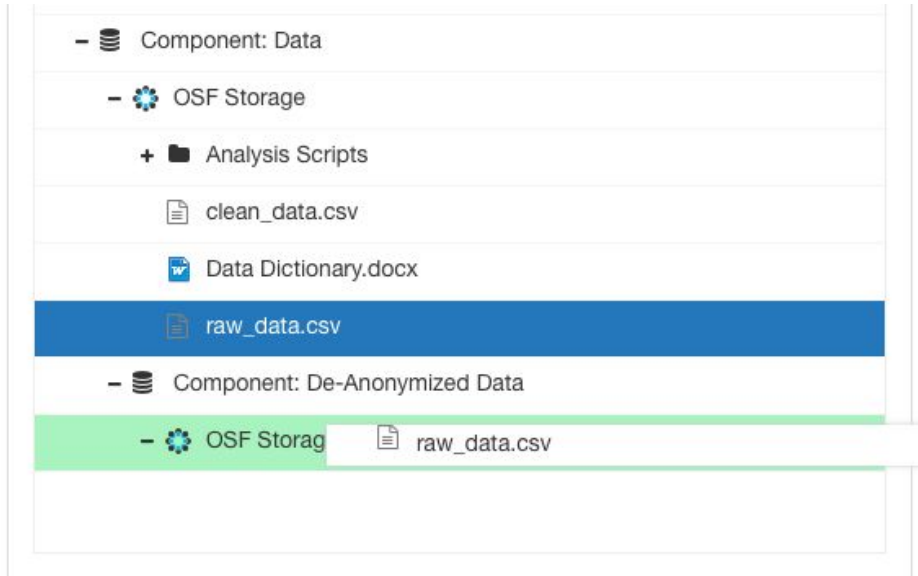
4. Archive + share your materials

Share your materials










- Where doesn't matter.
That you share matters.
- Get credit for your code, your data, your methods
- Increase the impact of your research



4. Share your work



The screenshot shows a file explorer interface with the following structure:

-  Component: Data
 -  OSF Storage
 - +  Analysis Scripts
 -  clean_data.csv
 -  Data Dictionary.docx
 -  raw_data.csv
-  Component: De-Anonymized Data
 -  OSF Storage
 -  raw_data.csv





Change privacy settings ×

Adjust your privacy settings by checking the boxes below.

Checked projects and components will be **public**.

Unchecked components will be **private**.


Select: [Make all public](#) | [Make all private](#)


- ☒ -  Demo Project
- ☒  Materials
- ☐  Data
- ☐  Background Literature




Cancel



Continue

4. Increasing discoverability




 Open Science Framework

My DashboardBrowse ▾Help ▾

 Courtney Soderberg


stereotype threat

Demo Project

PrivateMake Public 0 0

Contributors: [Courtney Soderberg](#), [Lauren Revere](#), [Casey Rollins](#)

Date Created: 2015-07-20 10:38 AM | Last Updated: 2015-07-21 03:40 PM

Category: Project 

Description: A project created to learn how to use the OSF.


OSF for Meetings

Open Science Framework

Browse - Support [Sign up](#) [Sign in](#)

OSF for Meetings


A **free poster and presentation sharing service** for academic meetings and conferences



For Conference Organizers

Register your event to broaden its impact. Events get a dedicated page, an easy submission process, and persistent links.

[Register](#)



For Conference Participants

Share your posters and presentations along with any supporting data and materials. Get persistent links and usage analytics.

[Upload](#)

[All meetings](#) [All submissions](#)

Only conferences with at least five submissions are displayed.

Search		
Name ^ v	Submissions ^ v	Accepting submissions ^ v
Society for Personality and Social Psychology 2016	484	Yes
Society for Personality and Social Psychology 2015	307	No
	29	

free poster +
presentation service

<https://osf.io/meetings>

OSF for Institutions

The screenshot shows the OSF website interface. At the top, the navigation bar includes the Open Science Framework logo, a search bar, and links for 'Browse', 'Support', 'Sign up', and 'Sign in'. Below the navigation bar, the 'CENTER FOR OPEN SCIENCE' logo is displayed, followed by a description: 'COS is a non-profit technology company providing free and open services to increase inclusivity and transparency of research. Find out more at cos.io.'

The main content area is titled 'All Projects' and features a filter bar. Below the filter bar, there is a table of projects and a sidebar with contributor information.

Collections	Name ^ v	Contributors	Modified ^ v
All Projects	COS Ambassadors	Bowman, Nose...	2 days ago
All Registrations	OSF for Institutions	Spitzer, Esposit...	2 days ago
	Badges to Acknowledge Op...	Blohowiak, Coh...	2 days ago
	OSF Public API Feedback	Geiger, Arslan +...	3 days ago
	Badges to Acknowledge Op...	Kidwell, Lazare...	6 days ago
	Reproducibility Project: Can..	Errington, Tan ...	6 days ago
	COS Reproducibility Works...	Soderberg	8 days ago

Contributors < 1/23 >

- Brian A. Nosek
- Sara Bowman
- Tim Errington
- Richard A. Klein

COS Ambassadors

Information Activity

Visibility : Public
Category: Project
Last Modified on: 2016-05-16 12:24 PM
A collection of materials and resources for COS ambassadors.

integration with local
research services

<https://osf.io/institutions/cos>

Transparency and Openness Promotion (TOP) Guidelines



<http://cos.io/top>

Citation standards

Data transparency

Analytic methods
(code) transparency

Research materials
transparency

Design and analysis
transparency

Preregistration
of studies

Preregistration
of analysis plans

Replication

Modular standards

Low barrier to entry

Discipline agnostic

<http://cos.io/top>

Citation standards

Data transparency

Analytic methods
(code) transparency

Research materials
transparency

Design and analysis
transparency

Preregistration
of studies

Preregistration
of analysis plans

Replication

Level 1 Disclose

Level 2 Require

Level 3 Verify

<http://cos.io/top>

LEVEL 1

LEVEL 2

LEVEL 3

Citation standards	Journal describes citation of data in guidelines to authors with clear rules and examples.	Article provides appropriate citation for data and materials used, consistent with journal's author guidelines.	Article is not published until appropriate citation for data and materials is provided that follows journal's author guidelines.
Data transparency	Article states whether data are available and, if so, where to access them.	Data must be posted to a trusted repository. Exceptions must be identified at article submission.	Data must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Analytic methods (code) transparency	Article states whether code is available and, if so, where to access them.	Code must be posted to a trusted repository. Exceptions must be identified at article submission.	Code must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Research materials transparency	Article states whether materials are available and, if so, where to access them.	Materials must be posted to a trusted repository. Exceptions must be identified at article submission.	Materials must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Design and analysis transparency	Journal articulates design transparency standards.	Journal requires adherence to design transparency standards for review and publication.	Journal requires and enforces adherence to design transparency standards for review and publication.
Preregistration of studies	Journal encourages preregistration of studies and provides link in article to preregistration if it exists.	Journal encourages preregistration of studies and provides link in article and certification of meeting preregistration badge requirements.	Journal requires preregistration of studies and provides link and badge in article to meeting requirements.
Preregistration of analysis plans	Journal encourages preanalysis plans and provides link in article to registered analysis plan if it exists.	Journal encourages preanalysis plans and provides link in article and certification of meeting registered analysis plan badge requirements.	Journal requires preregistration of studies with analysis plans and provides link and badge in article to meeting requirements.
Replication	Journal encourages submission of replication studies.	Journal encourages submission of replication studies and conducts blind review of results.	Journal uses Registered Reports as a submission option for replication studies with peer review before observing the study outcomes.

“

The policy of the __ is to publish papers where authors indicate whether the data, methods used in the analysis, and materials used to conduct the research will be made available to any researcher for purposes of reproducing the results or replicating the procedure.

- *Authors must, in acknowledgments or the first footnote, indicate if they will or will not make their data, analytic methods, and study materials available to other researchers.*
- *If an author agrees to make materials available, the author must specify where that material will be available.”*

<http://cos.io/top>

LEVEL 1

LEVEL 2

LEVEL 3

Citation standards	Journal describes citation of data in guidelines to authors with clear rules and examples.	Article provides appropriate citation for data and materials used, consistent with journal's author guidelines.	Article is not published until appropriate citation for data and materials is provided that follows journal's author guidelines.
Data transparency	Article states whether data are available and, if so, where to access them.	Data must be posted to a trusted repository. Exceptions must be identified at article submission.	Data must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Analytic methods (code) transparency	Article states whether code is available and, if so, where to access them.	Code must be posted to a trusted repository. Exceptions must be identified at article submission.	Code must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Research materials transparency	Article states whether materials are available and, if so, where to access them.	Materials must be posted to a trusted repository. Exceptions must be identified at article submission.	Materials must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Design and analysis transparency	Journal articulates design transparency standards.	Journal requires adherence to design transparency standards for review and publication.	Journal requires and enforces adherence to design transparency standards for review and publication.
Preregistration of studies	Journal encourages preregistration of studies and provides link in article to preregistration if it exists.	Journal encourages preregistration of studies and provides link in article and certification of meeting preregistration badge requirements.	Journal requires preregistration of studies and provides link and badge in article to meeting requirements.
Preregistration of analysis plans	Journal encourages preanalysis plans and provides link in article to registered analysis plan if it exists.	Journal encourages preanalysis plans and provides link in article and certification of meeting registered analysis plan badge requirements.	Journal requires preregistration of studies with analysis plans and provides link and badge in article to meeting requirements.
Replication	Journal encourages submission of replication studies.	Journal encourages submission of replication studies and conducts blind review of results.	Journal uses Registered Reports as a submission option for replication studies with peer review before observing the study outcomes.

“
The policy of the _____ is to publish papers only if the data used to conduct the research are clearly and precisely documented and are maximally available to any researcher for purposes of reproducing the results or replicating the procedure.

Details of:

- *What must be shared*
- *Legal and Ethical Exceptions – Disclosure at onset of review*
- *Using trusted repositories”*

<http://cos.io/top>

LEVEL 1

LEVEL 2

LEVEL 3

Citation standards	Journal describes citation of data in guidelines to authors with clear rules and examples.	Article provides appropriate citation for data and materials used, consistent with journal's author guidelines.	Article is not published until appropriate citation for data and materials is provided that follows journal's author guidelines.
Data transparency	Article states whether data are available and, if so, where to access them.	Data must be posted to a trusted repository. Exceptions must be identified at article submission.	Data must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Analytic methods (code) transparency	Article states whether code is available and, if so, where to access them.	Code must be posted to a trusted repository. Exceptions must be identified at article submission.	Code must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Research materials transparency	Article states whether materials are available and, if so, where to access them.	Materials must be posted to a trusted repository. Exceptions must be identified at article submission.	Materials must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Design and analysis transparency	Journal articulates design transparency standards.	Journal requires adherence to design transparency standards for review and publication.	Journal requires and enforces adherence to design transparency standards for review and publication.
Preregistration of studies	Journal encourages preregistration of studies and provides link in article to preregistration if it exists.	Journal encourages preregistration of studies and provides link in article and certification of meeting preregistration badge requirements.	Journal requires preregistration of studies and provides link and badge in article to meeting requirements.
Preregistration of analysis plans	Journal encourages preanalysis plans and provides link in article to registered analysis plan if it exists.	Journal encourages preanalysis plans and provides link in article and certification of meeting registered analysis plan badge requirements.	Journal requires preregistration of studies with analysis plans and provides link and badge in article to meeting requirements.
Replication	Journal encourages submission of replication studies.	Journal encourages submission of replication studies and conducts blind review of results.	Journal uses Registered Reports as a submission option for replication studies with peer review before observing the study outcomes.

“
... are maximally available
to any researcher for
purposes of reproducing the
results or replicating the
procedure.

*All materials supporting the
claims made by the author
must be made available to the
journal prior to
publication. The journal, or
an entity acting on behalf of
the journal, will verify that the
findings are replicable using
the author's data and
methods of analysis. Failure
to replicate at this stage may
result in the paper not being
published.”*

<http://cos.io/top>



ALFRED P. SLOAN
FOUNDATION



Society for
Personality
and Social
Psychology



Laura and John Arnold Foundation®



mozilla
Science Lab



nectar



Foundation for
Earth Science



an international forum for cell biology

539 Journals
59 Organizations



ASSOCIATION OF
RESEARCH LIBRARIES



Global Biological Standards Institute™



<http://cos.io/top>

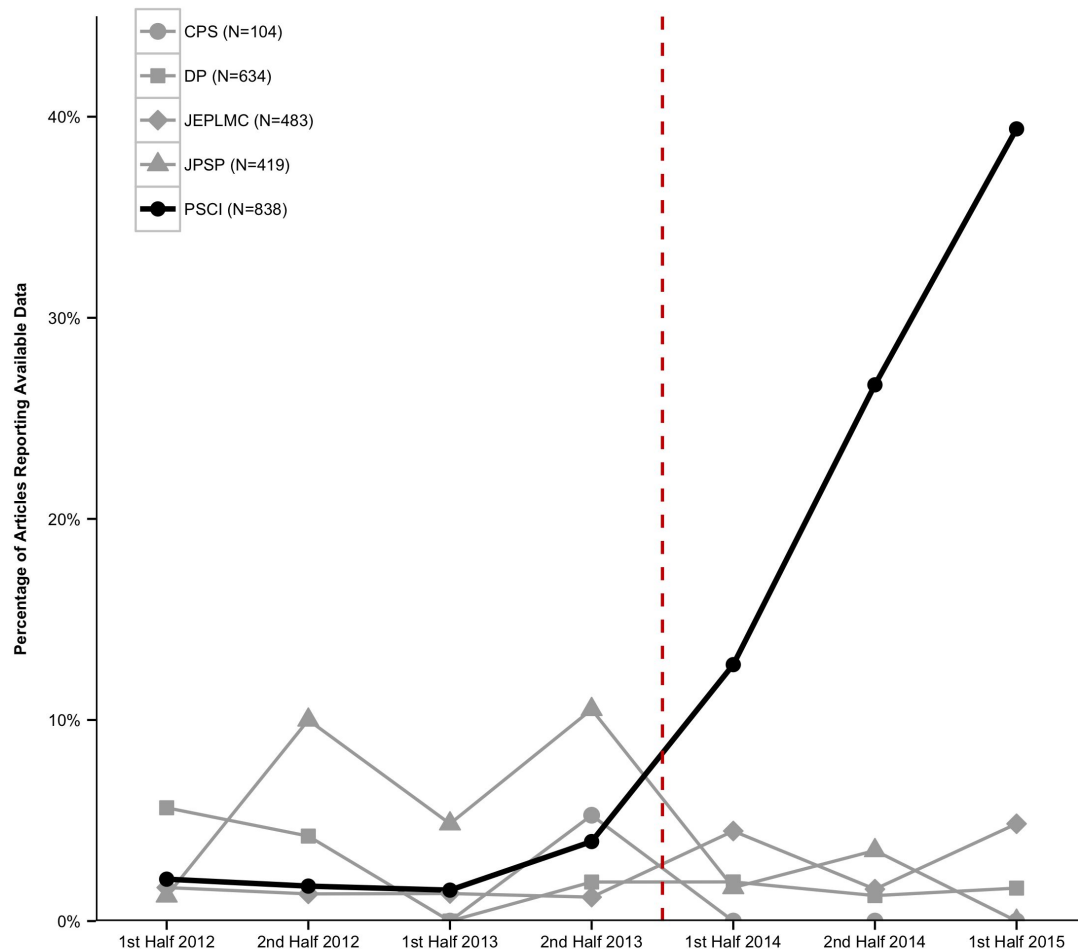
Badges for Open Practices



article signals open
behaviour

<https://osf.io/rfgdw>

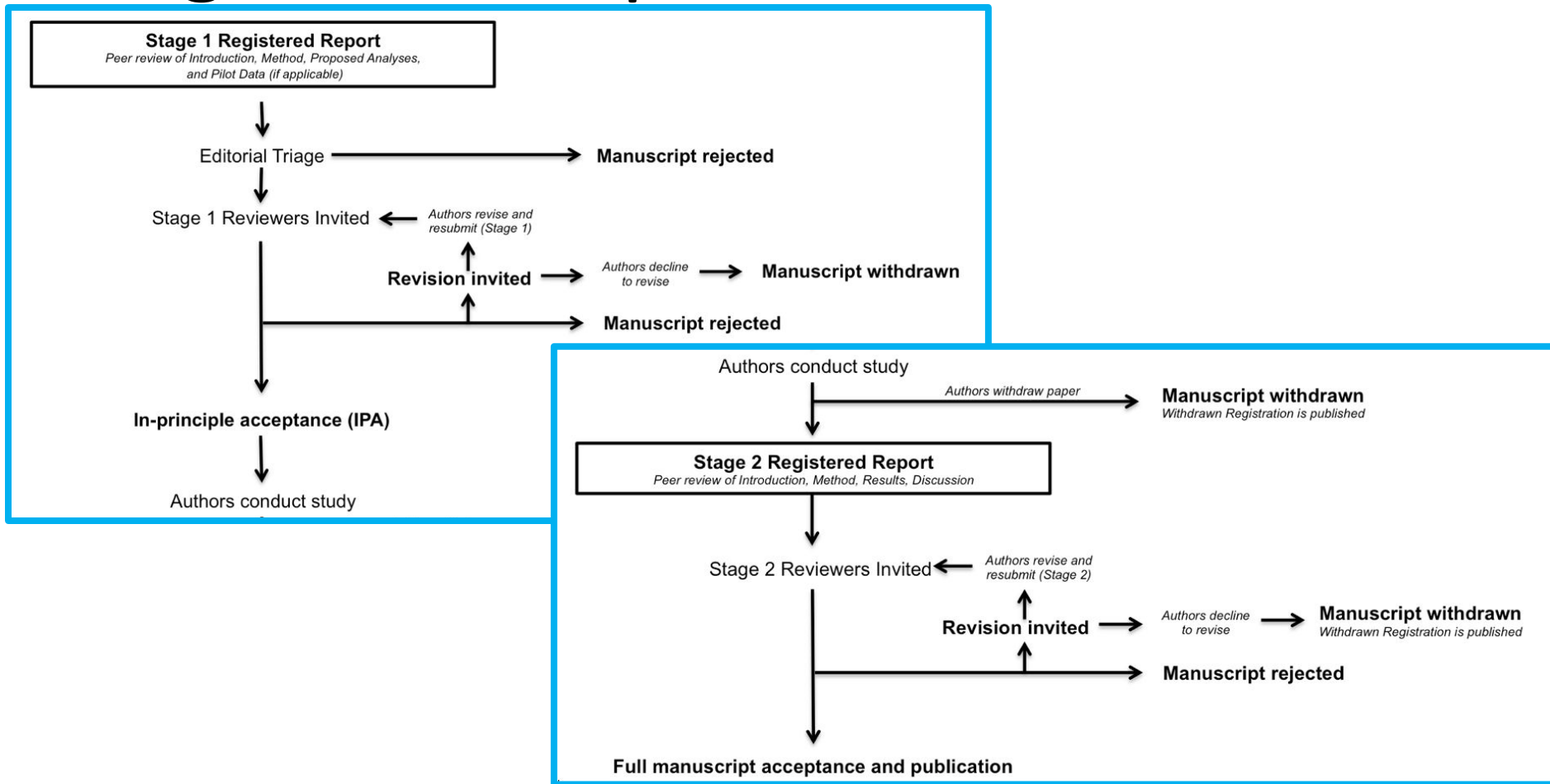
Case Study: Psychological Science



article signals open
behaviour

<https://osf.io/rfgdw>

Registered Reports



system to reward rigor
rather than findings

<https://osf.io/8mpji>

Reproducibility training

Statistical & Methodological Consulting

Scientists can improve the replicability of their own work through careful documentation, adherence to standards, and the use of open tools. We answer questions and provide training on open and reproducible tools, methodologies, and workflows. Some examples:

- ✓Using R
- ✓Conducting power analyses
- ✓Using the OSF
- ✓Learning Github
- ✓Conducting meta-analyses
- ✓Preregistering analysis plans

This service is provided in partnership with the [Berkeley Initiative for Transparency in the Social Sciences \(BITSS\)](#)



[Request a Consultation](#)



[Request a Workshop](#)



[Receive Updates
on our Services](#)

free stats + methods
training

http://cos.io/stats_consulting

How can you make your research reproducible?

1. Plan for reproducibility before you start

- Create a study plan – **Begin documentation at study inception**
- Set-up a reproducible project – **Centralize and organize your project management**
- Preregistration – **Preregister your study + analysis plan**
- Informative file naming - **The rules don't matter. That you have rules matters.**

2. Keep track of things

- Version control – **Track your changes**
- Documentation – **Document everything done by hand**

3. Contain bias

- Reporting – **Report transparently + completely**

4. Archive + share your materials

- Where doesn't matter. **That you share matters.**

How to learn more

- Organizing a project for reproducibility
 - [Reproducible Science Curriculum](#) by Jenny Bryan
 - <https://github.com/reproducible-science-curriculum/>
- Data management
 - [Data Management](#) from Software Carpentry by Orion Buske
 - <http://software-carpentry.org/v4/data/mgmt.html>
- Literate programming
 - [Literate Statistical Programming](#) by Roger Peng
 - <https://www.youtube.com/watch?v=YcJb1HBc-1Q>
- Version control
 - [Version Control](#) by Software Carpentry
 - <http://software-carpentry.org/v4/vc/>
- Sharing materials
 - Open Science Framework by Center for Open Science
 - <https://osf.io/>

Where to get help

reproducibility training

stats-consulting@cos.io

osf support

support@osf.io

work with COS

contact@cos.io

Q + A

- How might OSF fit into a workflow that uses USGS repositories for data publication (such as USGS ScienceBase, or NWIS)?
- How is OSF different from other workflow software such as Taverna, Kepler, etc.?
- How can OSF save me time? Generally takes large amounts of time to document.
- Can OSF document processing, analysis and results?
- What are some best practices for documenting scientific workflow for *new/exploratory methods* in data processing and analysis?